Fast Intra Prediction Algorithm in H.264/AVC

Fangwen Fu, Xinggang Lin, Lidong Xu
(Department of Electronic Engineering, Tsinghua University, Beijing China 100084)
Email: ffw98@mails.tsinghua.edu.cn

Abstract: Spatial intra prediction is introduced to achieve higher coding efficiency in H.264/AVC. Unfortunately, this comes at the cost in considerably increased complexity when RDO is enabled. In this paper, we propose a fast intra prediction algorithm including reducing the redundancy between the chroma mode decision and luma mode decision, improving the mode decision scheme and heuristically selecting block type. Experimental results show that the proposed algorithm can drastically reduce coding complexity with only negligible coding efficiency loss.

Key words: H.264/AVC, Intra Prediction, Rate Distortion Optimization (RDO), Mode Decision

1. Introduction

H.264 [1] is the ITU-T's newest video coding recommendation, which is also known as MPEG-4 Advanced Video Coding (AVC). H.264/AVC employs rate distortion optimization (RDO) techniques to achieve the best coding efficiency. During the implementation of RDO. H.264/AVC has to exhaustively compute rate and distortion of all the intra and inter prediction modes for encoding one macroblock(MB), which leads to drastic increase of coding complexity. Some fast prediction algorithms have ever been proposed to reduce the coding complexity [2][3]. In this paper, we propose an efficient algorithm to reduce the complexity of intra prediction when RDO is used.

In the luma intra prediction of H.264/AVC, two block types should be examined: 4x4 block and 16x16 block, called intra4x4 and intra16x16 respectively. Intra4x4 is based on predicting each 4x4 luma block separately and is well suited for coding parts of a picture with significant details. The intra16x16 performs the prediction of the whole 16x16 luma block and is more suited for coding very smooth areas of a picture [4]. In the chroma intra prediction there is only one block type: 8x8 block, named chroma8x8.

When the intra4x4 is used, each 4x4 block is predicted from spatially neighboring (above and left) samples (see Fig.1). For each 4x4 block, there are 9 prediction modes: DC prediction mode and eight directional prediction modes as shown in Fig.2. In the reference software JM7.3 [5], all the 9 modes are examined for a 4x4 block to find the optimal one with minimum cost, which is called *intra4x4 prediction routine*.

When the intra16x16 is adopted, four modes including vertical prediction, horizontal prediction, DC prediction and plane prediction are examined, which is called *intra16x16 prediction routine*.

Since chroma components (U and V) are usually smooth over large area [4], the prediction modes of the chroma samples in an MB are the same as those of intra16x16.

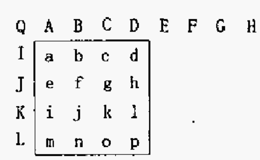

Fig.1. 4x4 bock and the neighboring samples

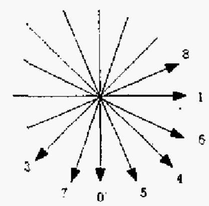

Fig.2. Eight prediciton modes for intra 4x4 prediction

In the past, researchers have proposed several algorithms to speedup the intra prediction module. An efficient intra prediction algorithm (EIP) [6] has been proposed for luma component based on the following main ideas: the examination can be early terminated if some conditions are satisfied. However, the conditions vary with video contents and it is difficult to find good adaptive conditions. A fast mode decision algorithm [7] for intra prediction was also developed, in which Pan pointed out that the pixels along the direction of local edge are normally of similar values for both luma and chroma components. In Pan's algorithm, the edge information of the whole picture is firstly produced with the Sobel operator [8]. Then the edge direction histograms are calculated for 4x4 luma, 16x16 luma and 8x8 chroma blocks, respectively. In intra prediction for

0-7803-8406-7/04/$20.00 © 2004 IEEE.

4x4 blocks, the histogram cell with the maximum amplitude and the two adjacent cells are selected together with DC prediction mode as candidates for the best 4x4 prediction mode. In intra prediction for 16x16 blocks, the histogram cell with maximum amplitude and DC prediction mode are considered as candidates of the best 16x16 prediction mode. In chroma prediction the histogram cell with maximum amplitude in U component, the one in V component and DC mode are selected as candidates for the best chroma prediction mode. Although Pan's algorithm has reduced much complexity of intra prediction, it still needs much computation. In this paper, we proposed a more efficient algorithm by modifying prediction routine, improving mode decision scheme and adopting a heuristic block type selection method.

The paper is organized as follows. The proposed algorithm is explained in section 2, followed by simulation results in Section 3. In Section 4 the conclusion is addressed.

2. Proposed Algorithms

2.1 Modified Prediction Routine

As studied, the intra prediction procedure for YUV components in the reference software of H.264/AVC can be depicted as follows:

1. Generate an 8x8 predicted chroma block according to a mode.
2. Perform the *intra16x16 prediction routine* to select the best 16x16 mode, code the chroma components with the given mode, and then calculate the rate distortion of the MB for YUV components $RDCost_{16x16}$;
3. Perform the *intra4x4 prediction routine* for 16 4x4 blocks to select the 16 best 4x4 modes, code the chroma components with the given mode, and then calculate the rate distortion of the MB for YUV components $RDCost_{4x4}$.
4. If $RDCost_{16x16} > RDCost_{4x4}$, the block type 4x4 is selected, otherwise, the 16x16 block type is selected in the given chroma mode. And the minimum cost is saved as $RDCost$.
5. Repeat step 1 to 4 for all four chroma intra prediction modes, and choose the one with minimum $RDCost$.

From the above procedure, it is obvious that chroma components should be coded twice for each prediction mode and *Intra16x16 prediction routine* and *intra4x4 prediction routine* should be executed four times respectively when an MB is encoded with intra prediction. It can also be seen that the rate distortion of YUV components should be computed eight times. Thus, the chroma mode decision procedure is highly correlated to that of luma. However, the two procedures are decorrelated with the following steps in the proposed algorithm.

1. Perform *chroma intra8x8 prediction routine*
   a. Generate an 8x8 prediction block according to a mode for UV;
   b. Apply transform operation and quantization on prediction error, and then apply dequantization and inverse transform to obtain decoded block for UV, respectively.
   c. Calculate the distortion CD_{8x8} between the original 8x8 block and the decoded block for U and V components;
   d. Calculate the bitrate $CRate_{8x8}$ for coding the U and V transform coefficients and prediction modes, then compute $Cost_{8x8} = CD_{8x8} + \lambda(Q_p)CRate_{8x8}$,where $\lambda(Q_p)$ is an exponential function of the quatization factor;
   e. Repeat step a to d for all the 4 chroma prediction modes, and then choose the one with the minimum cost.
2. Perform *intra16x16 prediction routine*, and calculate $RDCost_{16x16}$.
3. Perform *intra4x4 prediction routine*, and calculate $RDCost_{4x4}$
4. If $RDCost_{16x16} > RDCost_{4x4}$, intra4x4 is selected, otherwise, intra16x16 is selected.

In this modified routine, each chroma prediction mode is only examined once, and the intra16x16 prediction routine and intra4x4 prediction routine are also performed once, respectively. The experimental results show that the modified routine can drastically reduce the computational complexity of intra prediction.

2.2 Improved Mode Decision Algorithm

In Pan's algorithm, the Sobel operator is applied to all pixels of the whole picture. But in the experiment, we observed that the mode decision normally depends on the edge information between the left, above blocks and current block. So it's not necessary to compute the edge information of all pixels in the whole picture. Instead, we only calculate the edge information of pixels in the boundaries of 4x4 blocks for luma and of 8x8 blocks for chroma. Then the edge direction histograms are computed for intra4x4, intra16x16 and chroma8x8 blocks similar to Pan's. For 4x4 luma blocks, the histogram cell with the maximum amplitude and two adjacent cells are selected together with DC prediction mode as the candidate modes. Because of the correlation between spatially adjacent blocks, the most probable mode generated from the final selected modes of above and left blocks is also considered as a candidate. For 16x16 luma blocks and 8x8 chroma blocks, only the histogram cell with maximum amplitude and DC mode are selected as candidates.

2.3 Block Type Selection Method

In the experiment, it's observed that the luma coding block type of the current MB is highly correlated to that of its neighboring MBs. It is also seen that the correlation decreases when quantization parameter Q_p increases. When the current MB is encoded, the coding block types and costs of the left and above MBs have been known. Similar to intra4x4 luma coding scheme, a most probable block type for the current MB can be obtained. If the most probable block type is intra4x4, then intra4x4 is first examined. In this case, if the cost of intra4x4 is very large ($> T_1$), which means that this MB is chaotic, it is very likely that intra16x16 will not be chosen. If the most probable block type is intra16x16, then intra16x16 is first examined. In this case if the cost of intra16x16 is small enough ($< T_2$), which means this MB is smooth, it is very likely that intra4x4 should not be examined.

The proposed block type selection method for luma component and the threshold updating of T_1 and T_2 are given as follows:

1. If the MB is at the top or left of the frame, intra4x4 and intra16x16 are both examined and the type with minimum cost is selected as the optimal one, then go to step 4. Otherwise go to step 2.
2. Take the block type of the above or left MBs with minimum cost as the most probable block type and examine it. If the most probable block type is intra4x4 and the cost of current MB is larger than T_1, choose this block type as the best one, go to step 4. Otherwise, if the most probable block type is intra16x16 and the cost is smaller than T_2, choose this block type as the best one, go to step 4. Otherwise, go to step 3.
3. Examine the other block type, and choose the one with minimum cost as the best one.
4. If there are more MBs, consider the next MB and go to step 1.

In step 2, T_1, T_2 is updated as follows:

If block types of the above and left MBs are both intra4x4, then the most probable type is intra4x4 and

$$T_1 = \min\{RDCost_{above}, RDCost_{left}\} + 15 * \lambda(Q_p).$$

If block types of the above and left MBs are both intra16x16, then the most probable type is intra16x16 and

$$T_2 = \max\{RDCost_{above}, RDCost_{left}\}.$$

If block types of the above and left MBs are different, then the most probable type is the one with minimum cost. If the most probable type is intra4x4, then

$$T_1 = RDCost_{intra4x4} + 15 * \lambda(Q_p).$$

otherewise, $T_2 = RDCost_{intra16x16}$.

3. Experimental results

The proposed algorithm is tested on three CIF sequences, Container, News and Stefan, and compared with Pan's algorithm and the full search (FS). For each sequence, 298 frames are encoded with I-frame only and with various Q_p. The experiments were implemented on JM7.3 provided by JVT and the platform is Intel PIII, 730MHz and 256MB memory. The test conditions are: 1.Hadamard transform is used; 2. RD optimization is enabled; 3. CABAC is enabled.

The test results are shown in Table 1-3. In the method column A, B and C represent our proposed algorithm, Pan's algorithm and full search algorithm, respectively, in the-method column. From the tables, we can see that our proposed algorithm can reduce computational complexity by about 4 times as FS with little decrease in PSNR and a slight increase in bit rate. The proposed algorithm, compared with Pan's, can reduce computation time by about 2 times and also outperform the Pan's in PSNR of U and V components only with a little decrease in PSNR of Y components.

4. Conclusion

In this paper, an efficient intra prediction algorithm is proposed for RDO in H.264/AVC. There are three major contributions to the reduction of the complexity of intra prediction: 1) reducing the correlation between chroma mode decision and luma mode decision; 2) improving mode decision based on Pan's; 3) heuristically selecting block types for luma. Experimental results show that the proposed algorithm can achieve considerable computation reduction at the price of negligible coding efficiency loss compared with FS. The proposed algorithm is also better than Pan's in complexity reduction.

References

[1] *Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14496-10 AVC)* in JVT of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, 2003

[2] Chen Z.B., Zhou P., He Y., *Fast IntegerPel and Fractional Pel Motion Estimation for JVT*, JVT-F017, 6th Meeting, Awaji Island, Japan, December 5-13, 2002

[3] Ahmad A., Khan N., Masud S., and Maud M.A., *Efficient Block Size Selection in H.264 video coding standard*, Electronics Letters, 40 (1), pp. 19-21, 2004

[4] Wiegand T., Sullivan G.J., Bjontegaard G., and Luthra A., *Overview of the H.264/AVC video coding standard*, IEEE Trans. Circuits Syst. Video Technol., 13 (7), pp. 560-576, July 2003

[5] Joint Video Team (JVT), *reference software JM7.3*, ftp://ftp.imtc-files.org/jvt-experts

[6] Meng B., Au O.C., Wong C.W. and Lam H.K., *Efficient Intra-Prediction Algorithm in H.264*, Proceedings of International Conference on Image Processing, 2003, pp. 837-840

[7] Pan F., Lin X., Susanto R., Lim K.P., Li Z.G., Feng G.N., Wu D. J. and Wu S., *Fast Mode Decision Algorithm for Intra Prediction in JVT*, G013, 64th MPEG and 7th JVT meeting, Pattaya, March 2003

[8] Rafael C.G., Richard E.W., "*Digital Image Processing,*" Prentice Hall, 2002

Table 1: Results of 'Container'

Qp	Method	Y (dB)	U (dB)	V (dB)	Time (s)	Bit rate (Kbps)
24	A	39.84	44.38	44.64	284	5358.2
	B	39. 85	44.35	44.56	583	5335.7
	C	39.92	44.45	44.66	1312	5291.3
32	A	34. 30	40.46	40.63	201	1642.9
	B	34. 31	40.41	40.60	512	1647.1
	C	34. 35	40.47	40.67	1089	1600.0
40	A	29.33	37.68	37.58	182	678.0
	B	29.36	37.71	37.46	489	694.6
	C	29.36	37.60	37.62	971	652.0

Table 2: Results of 'News'

Qp	Method	Y (dB)	U (dB)	V (dB)	Time (s)	Bit rate (Kbps)
24	A	41.57	43. 84	44. 65	239	2890.6
	B	41. 57	43.80	44.65	589	2893.0
	C	41. 63	43.87	44.66	1219	2817.5
32	A	35. 98	40.06	40.90	197	1425.4
	B	36. 01	40.00	40.75	500	1442.4
	C	36. 02	40.06	40.85	1054	1380.6
40	A	30.45	36.81	37.82	183	670.8
	B	30.53	36.85	37.69	471	683.3
	C	30.53	36.76	37.92	964	638.7

Table 3: Results of 'Stefan'

Qp	Method	Y (dB)	U (dB)	V (dB)	Time (s)	Bit rate (Kbps)
24	A	39. 95	41. 64	41. 69	303	5979. 8
	B	39. 93	41.61	41.65	647	5961.3
	C	40. 07	41.67	41.72	1480	5878.4
32	A	33. 28	37.09	37.04	250	3131.7
	B	33. 27	37.07	37.02	560	3132.2
	C	33. 38	37.07	37.04	1224	3056.8
40	A	27.12	33.86	33.66	208	1392.0
	B	27.13	33.88	33.68	503	1406.1
	C	27.20	33.79	33.56	1047	1345.1